GST SEQUENCES FROM SOYBEAN AND THEIR USE IN THE PRODUCTION OF HERBICIDE RESISTANT PLANTS

The present invention relates *inter alia*, to Glutathione-*S*-transferase (GST) sequences and their use in methods for the production of herbicide resistant plants. In particular the polynucleotides according to the invention may be used in methods for the production of plants which are resistant to herbicides comprising fomesafen and/or acifluorfen.

Plants which are substantially "tolerant" to a herbicide when they are subjected to it provide a dose/response curve which is shifted to the right when compared with that provided by similarly subjected non tolerant like plants. Such dose/response curves have "dose" plotted on the x-axis and "percentage kill", "herbicidal effect" etc. plotted on the y-axis. Tolerant plants will typically require at least twice as much herbicide as non tolerant like plants in order to produce a given herbicidal effect. Plants which are substantially "resistant" to the herbicide exhibit few, if any, necrotic, lytic, chlorotic or other lesions when subjected to the herbicide at concentrations and rates which are typically employed by the agrochemical community to kill weeds in the field. Hereinafter the words (i) "tolerant" and (ii) "resistant" when used individually mean "tolerant and/or resistant".

Herbicide resistant plants are already available within the art for example, ROUNDUP READY™ Soya which is resistant to herbicides having as a site of action the enzyme 5-enolpyruvylshikimate-3-phosphate synthase, such as those agrochemicals containing glyphosate. One of the advantages of these plants is that the farmer can apply the herbicides to fields containing the resistant crop plants and weeds using "over-the-top application", to kill the weeds.

Other examples of products for use in methods for the production of herbicide resistant plants are provided in International Patent Application Publication Number WO 93/01294 and WO99/14337. Here the resistance is achieved by inserting into the plant a polynucleotide which provides for the production of a glutathione-S-transferase (GST) enzyme which is involved with the detoxification of the herbicide. Glutathione-S-transferase enzymes have been shown to exist in various organisms such as bacteria, fungi, yeast, plants, mammals

and fish and may exist as homo or heterodimers with subunits typically between 24 and 30 KDa. It has been shown that herbicide detoxification is achieved by the conjugation of the herbicide with the free thiol glutathione (GSH), a tripeptide (gamma-glutamyl-cysteinyl-glycine) within the plant (Cole D.J. 1994 Pesticide Science. 42 pp209-222). Such conjugation is catalysed by GST. Detoxification of herbicides has also been shown to occur following the conjugation of the herbicide with homoglutathione, which is the predominant thiol in some leguminous species. Homoglutathione (hGSH) is also a tripeptide (gamma-glutamyl-cysteinyl-Beta-alanine) but differs from GSH by the addition of Beta-alanine instead of a glycine to the gamma-glutamyl-cysteinyl part.

Thus, the present invention seeks to provide *inter alia*, novel polynucleotides which encode proteins which can be used in methods of providing plants with high levels of resistance to a herbicide which comprises fomesafen and/or acifluorfen.

According to the present invention there is provided a Glutathione-S-transferase (GST) comprising the amino acid sequence depicted as SEQ ID No. 10 or a variant GST having at least 80% identity therewith with the *proviso* that said variant GST does not comprise the amino acid sequence depicted as SEQ ID No. 36. (which corresponds to the sequence of clone SE3.03B09 listed as SEQ ID No. 8 in International Patent Application Publication Number WO00/18936). In a further embodiment of the present invention said variant GST has at least 85% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 90% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 91% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 92% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 93% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 94% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 95% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at

least 96% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 97% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 98% identity to the sequence depicted as SEQ ID No. 10. In a still further embodiment of the present invention said variant GST has at least 99% identity to the sequence depicted as SEQ ID No. 10. The present invention still further provides a GST comprising the sequence depicted as SEQ ID No. 7, 8 or 9 or a GST variant having at least 80% identity therewith with the *proviso* that said variant does not encode the sequence depicted as SEQ ID No. 37 (which corresponds to the sequence of clone SS1.PK0014.A1 listed as SEQ ID No. 22 in International Patent Application Publication Number WO00/18936). The percentage of sequence identity for proteins is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the amino acid sequence in the comparison window may comprise additions or deletions (i.e. gaps) as compared to the initial reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of match positions, dividing the number of match positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. When calculating the percentage sequence identity the sequences may be aligned allowing for up to 3 gaps with the *proviso* that in respect of the gaps, a total of not more than 15 amino acid residues is affected. Optimal alignment of sequences for comparison may also be conducted by computerised implementations of known algorithms. In a particular embodiment of the present invention the sequence identity is calculated using the FASTA version 3 algorithm which uses the method of Pearson and Lipman (Lipman, D.J. and Pearson, W.R. (1985) Rapid and Sensitive Protein Similarity Searches and Science. 227:1435-1441 and Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. PNAS. 85:2444-2448) to search for similarities between the reference sequence (also termed the query sequence) and any group of sequences (termed further sequences). Methods also exist in the art which enable

the percentage sequence identity between polynucleotide sequences to be calculated.

The protein may differ from the basic GST protein sequence (such as SEQ ID No. 10) by conservative or non-conservative amino acid substitutions. A conservative substitution is to be understood to mean that the amino acid is replaced with an amino acid with broadly similar chemical properties. In particular conservative substitutions may be made between amino acids within the following groups:

(i) Alanine and Glycine;

(ii) Serine and Threonine;

(ii) Glutamic acid and Aspartic acid;

(iii) Arginine and Lysine;

(iv) Asparagine and Glutamine;

(v) Isoleucine and Leucine,

(vi) Valine and Methionine;

(vii) Phenylalanine and Tryptophan.

In general, more conservative than non-conservative substitutions will be possible without destroying the GST properties of the proteins. Suitable variant proteins in accordance with the present invention may be determined by testing the GST properties of the protein using routine methods which are well known to the person skilled in the art. Such variant proteins may also be synthesised chemically using standard techniques.

The present invention further provides a GST or variant as described above wherein said GST or variant is capable of conferring resistance and/or tolerance to a herbicide which comprises fomesafen and/or acifluorfen. The herbicide may also comprise other Diphenyl ethers or Sulphonylureas such as Chlorimuron Ethyl and/or Chloroacetanilides such as acetochlor.

The present invention still further provides a polynucleotide comprising a region which encodes a GST or a variant GST as described above. In a particular embodiment of the present invention said polynucleotide comprises the sequence depicted as SEQ ID No. 14.

The present invention still further provides a polynucleotide sequence which is the complement of one which binds to a polynucleotide as described above at a

temperature of between 60°C and 65°C in 0.3 strength citrate buffered saline containing 0.1% SDS followed by rinsing at the same temperature with 0.3 strength citrate buffered saline containing 0.1% SDS wherein said polynucleotide sequence still encodes a functional GST with the *proviso* that said polynucleotide sequence is not the sequence depicted as SEQ ID No. 38. This sequence corresponds to the sequence of clone SE3.03B09 listed as SEQ ID No. 7 in International Patent Application Publication Number WO00/18936. The aforementioned hybridisation conditions are well known to the person skilled in the art as high stringency conditions. As an alternative to the protocol mentioned above the skilled person may identify a polynucleotide sequence encoding a GST variant according to the present invention under the following conditions: hybridisation at a temperature of about 65°C in a solution containing 6 X SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 X SSC and 0.1% SDS.

The present invention still further provides a polynucleotide sequence as described above which is not a sequence selected from the group of sequences listed under accession numbers: AF004358; AC000348; AF051214; AF051238 and P32110.

The present invention still further provides a polynucleotide comprising the sequence depicted as SEQ ID No. 11, 12 or 13 or a polynucleotide sequence which is the complement of one which hybridises to any one of SEQ ID Nos. 11, 12 or 13 under stringent conditions with the *proviso* that said polynucleotide sequence is not a sequence selected from the group depicted as SEQ ID Nos. 39 to 43. These sequences correspond to the sequence of clones SS1.PK0014.A1; SES8W.PK0028.C6; SR1.PK0011.D6; SS1.PK0020.B10; SSM.PK0067.G5 listed as SEQ ID Nos. 21, 13, 15, 23 and 25 respectively in International Patent Application Publication Number WO00/18936.

The present invention further provides a polynucleotide sequence as described in the preceding paragraph wherein said polynucleotide sequence or the amino acid sequence which it encodes, is not a sequence selected from the group of sequences listed under accession number; P32110, U20809, Q03663, P32111, P46421, AJ000923, AF004358, AC000348, AF051214 or AF051238. These accession numbers relate to GST and putative GST sequences in the database (and

also depicted in the sequence listing) which exhibit a low % identity to the sequences according to the present invention.

In a further aspect of the present invention there is provided a protein comprising the amino acid sequence depicted as SEQ ID No. 1 or a protein variant having at least about 70% identity therewith wherein the said protein or variant is capable of catalysing the addition of Beta-alanine onto gamma glutamylcysteine. In a further embodiment of the present invention said protein variant has at least 75% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 80% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 85% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 90% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 95% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 96% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 97% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 98% identity with the sequence depicted as SEQ ID No. 1. In a still further embodiment of the present invention said protein variant has at least 99% identity with the sequence depicted as SEQ ID No. 1. The sequence identity may be calculated and the protein variants created/identified in a manner analogous to that described above.

The present invention still further provides a protein variant as described in the preceding paragraph having a Km for Beta-alanine which is less than the said variants Km for glycine when calculated using the same method. In a further embodiment of the present invention said protein variant has a Km for Beta-alanine which is less than or equal to about 1mM and a Km for glycine which is higher than 1mM when calculated using the same method. In a still further embodiment of the present invention said protein variant has a Km for Beta-alanine which is less than or equal to about 0.8mM and a Km for glycine which is higher than 0.8mM when calculated using the same method. The Km or "Michaelis-Menten Constant" is a

kinetic parameter which indicates the substrate concentration at which the initial velocity of the reaction (V_0) is half maximal. Methods for the calculation of Km are well known to the person skilled in the art.

The present invention further provides a protein variant as described above which has a Vmax for Beta-alanine which is greater than the said variants Vmax for glycine when calculated using the same method. Vmax is also know the skilled person as the maximum rate of enzyme reaction i.e. the point at which the enzyme becomes saturated with substrate.

The present invention further provides a protein variant as described above comprises a sequence which contains at least one of the amino acid sequence regions depicted in the group depicted as SEQ ID No. 2 (KKIQQELAKP); SEQ ID No. 3 (CFAGLWSL); SEQ ID No. 4 (VMKPQREGGGNNIYG) and SEQ ID No. 5 (AAYILMQRIFP). In a still further embodiment of the present invention said protein variant comprises a sequence which contains at least two amino acid sequence regions selected from the group depicted as SEQ ID No. 2, 3, 4 or 5. In a still further embodiment of the present invention said protein variant comprises a sequence which contains all of the amino acid sequence regions depicted as SEQ ID No. 2, 3, 4 or 5

The present invention still further provides a polynucleotide comprising a region encoding the protein or protein variant as described above. In a further embodiment of the present invention said polynucleotide comprises the sequence depicted as SEQ ID No. 6.

The present invention still further provides a polynucleotide sequence which is the complement of one which hybridises to the polynucleotide depicted as SEQ ID No. 6 under stringent conditions wherein said polynucleotide sequence still encodes a protein which is capable of catalysing the addition of Beta-alanine onto gamma glutamylcysteine. Such hybridisation conditions are described above.

The present invention still further provides a polynucleotide comprising a first region comprising a polynucleotide encoding a GST or variant GST according to the present invention and a second region comprising a polynucleotide encoding a protein or protein variant according to the present invention. In a further embodiment of the present invention said first region comprises a polynucleotide encoding the amino acid sequence depicted as SEQ ID No. 10 and said second

region comprises a polynucleotide encoding the amino acid depicted as SEQ ID No. 1. The regions may be separated by a region which provides for a self processing polypeptide which is capable of separating the proteins such as the self processing polypeptide described in US5,846,767 or any similarly functioning element. Alternatively the regions may be separated by a sequence which acts as a target site for an external element which is capable of separating the protein sequences. Alternatively the polynucleotide may provide for a polyprotein which comprises a plurality of protein functions such as a GST and a homoglutathione synthetase according to the invention. In a further embodiment of the present invention the proteins of the polyprotein may be arranged in tandem. In a still further embodiment of the present invention the polyprotein comprises a plurality of protein functions which are separated by linker sequences.

The present invention still further provides a method of evolving a polynucleotide which encodes a GST protein which is capable of conferring resistance to a herbicide which comprises fomesafen and/or acifluorfen or a protein which is capable of catalysing the addition of Beta-alanine onto gamma glutamylcysteine (i.e. having such catalytic properties) comprising: (a) providing a population of variants of said polynucleotide and further polynucleotides which encode further proteins, at least one of which is in cell free form; and (b) shuffling said variants and further polynucleotides to form recombinant polynucleotides; and (c) selecting or screening for recombinant polynucleotides which have evolved towards encoding a protein having the said GST or said catalytic properties; and (d) repeating steps (b) and (c) with the recombinant polynucleotides according to step (c) until an evolved polynucleotide which encodes a protein having insecticidal properties has been acquired wherein said population of variants in part (a) contains at least a polynucleotide encoding a GST or variant GST protein or a protein or protein variant as described above.

The present invention still further provides a method as described above wherein said population of variants in part (a) contains at least a polynucleotide encoding the protein depicted as SEQ ID No. 10 and said further polynucleotides in part (a) encode a protein depicted as SEQ ID No. 1. The methods for evolving a polynucleotide as described above are well known to the person skilled in the art and are described *inter alia*, in US Patent No. 5,811,238.

The present invention still further provides a polynucleotide obtainable or obtained by the methods described above and a protein encoded by any such polynucleotide.

The present invention still further provides a DNA construct comprising in sequence a plant operable promoter operably linked to a polynucleotide according to the present invention operably linked to a transcription termination region. In a further embodiment of the present invention the DNA construct further comprises a region or a plurality of regions which provide for the targeting of the protein product or products to a particular location or locations. The DNA construct may further comprise a region which provides for the production of a protein which acts as a selectable marker. The selectable marker may, in particular, confer resistance to kanamycin; hygromycin or gentamycin. Further suitable selectable markers include genes which confer resistance to other herbicides such as glyphosate based herbicides or resistance to toxins such as eutypine. Other forms of selection are also available such as hormone based selection systems such as the Multi Auto Transformation (MAT) system of Hiroyrasu Ebinuma *et al.* 1997. PNAS Vol. 94 pp2117-2121; visual selection systems which use the known green fluorescence protein, β glucoronidase and any other selection system such as mannose isomerase, xylose isomerase and 2-deoxyglucose (2-DOG). The plant operable promoter of the DNA construct may be selected from the group consisting of *Agrobacterium rhizogenes* RolD; RolD/Fd; potato protease inhibitor II; CaMV35S; CamV35S double enhanced; FMV35S; NOS; OCS; Patatin; E9; alcA/alcR switch; GST switch; RMS switch; oleosin; ribulose bisphosphate carboxylase-oxygenase small sub-unit promoter. Terminators which can be used in the constructs according to the present invention include Nos, proteinase inhibitor II and the terminator of a gene of alpha-tubulin (EP-A 652,286). It is equally possible to use, in association with the promoter regulation sequence, other regulation sequences which are situated between the promoter and the sequence encoding the protein according to the present invention, such as transcriptional or translational enhancers, for example, tobacco etch virus (TEV) translation activator described in International Patent application, PCT publication number WO87/07644 or the Glucanase II leader sequence. The polynucleotide encoding the protein according to the invention may also be codon-optimised, or otherwise altered to enhance for

example, transcription once it is incorporated into plant material. Such codon optimisation may also be used to alter the predicted secondary structure of the RNA transcript produced in any transformed cell, or to destroy cryptic RNA instability elements present in the unaltered transcript, thereby increasing the stability and/or availability of the transcript in the transformed cell (Abler and Green. 1996. Plant Molecular Biology (32) pp63-78).

The present invention still further provides a method of providing plants which are resistant and/or tolerant to an agrochemical comprising: (a) inserting into the genome of plant material a polynucleotide or a polynucleotide sequence which provides for a GST or variant GST as described above or a DNA construct as described above; and (b) regenerating plants or plant parts therefrom; and (c) applying to said plants or plant parts an amount of said agrochemical which is phytotoxic to control-like plants and selecting those plants or plant parts which are resistant to said agrochemical. In a further embodiment of the present invention the polynucleotide inserted into plant material in accordance with the method of the preceding sentence encodes an amino acid sequence depicted as SEQ ID No. 10. The present invention still further provides a method of providing plants which are resistant and/or tolerant to an agrochemical comprising: (a) inserting into the genome of plant material from a plant which provides for the production of a functional GST, a polynucleotide encoding a protein or variant which is capable of catalysing the addition of Beta-alanine onto gamma glutamylcysteine as described above or a DNA construct as described above; and (b) regenerating plants or plant parts therefrom; and (c) applying to said plants or plant parts an amount of said agrochemical which is phytotoxic to control like plants and selecting those plants or plant parts which are resistant to said agrochemical.

The present invention still further provides methods as described in the preceding paragraph wherein said agrochemical comprises fomesafen and/or acifluorfen. The polynucleotide/DNA construct may be incorporated into the genome of plant material in accordance with the present invention by plant transformation techniques which are well known to the person skilled in the art. Such techniques include but are not limited to particle mediated biolistic transformation, *Agrobacterium*-mediated transformation, protoplast transformation (optionally in the presence of polyethylene glycols); sonication of plant tissues,

cells or protoplasts in a medium comprising the polynucleotide or vector; micro-insertion of the polynucleotide or vector into totipotent plant material (optionally employing the known silicon carbide "whiskers" technique), electroporation and the like. Techniques which are specifically optimised for a particular crop may also be employed for the purposes of transforming a particular crop in accordance with the present invention. For example, when transforming soybean it may be preferable to use a method based on one described by Christou *et al* (1990), Dan *et al* (1999), Williams *et al* (2000) and Hinchee *et al* (1999). The transformation method *per se* is not germane to the present invention and the skilled person may employ any functional method appropriate to the target crop.

The present invention still further provides herbicide resistant plants or plant parts obtained according to the methods described above. Such plants or plant parts may be selected from the group consisting of: melons; mangoes; soybean; cotton; tobacco; sugarbeet; oilseed rape; canola; flax; sunflower; potato; tomato; alfalfa; lettuce; maize; wheat; sorghum; rye; bananas; barley; oat; turf grass; forage grass; sugar cane; pea; field bean; rice; pine; poplar; apple; peaches; grape; strawberries; carrot; lettuce; cabbage; onion; citrus; cereal; nut plants or other horticultural crops. In a particular embodiment of the present invention said plants or plant parts are soybean (*Glycine sp.*) plants or plant parts.

The present invention still further provides the use of a polynucleotide encoding a GST or variant GST or a DNA construct as described above in a method of producing plants which are resistant and/or tolerant to a herbicide comprising fomesafen and/or acifluorfen.

The present invention still further provides a method of providing a plant with a further desired agronomic trait comprising: (a) inserting into the genome of plant material from a plant or plant part as described above a polynucleotide which provides for the desired agronomic trait; and regenerating plants or plant parts from said material; or (a) crossing a first plant or plant part as described above with a second plant which provides for said desired agronomic trait; and (b) selecting those resultant plants which contain said further desired agronomic trait. It will be appreciated that the plant resultant from such method will contain the characteristics of the GSTs and/or homoglutathione synthetase according to the present invention as described above along with the characteristics associated with said further

agronomic trait. In a further embodiment of the present invention said further desired agronomic trait provides resistance to a herbicide which comprises glyphosate or a salt thereof. The further desired agronomic trait may also be selected from the group consisting of: further herbicide resistance; insect resistance; nematode resistance; stress tolerance; altered yield; altered nutritional value; altered quality or any other desirable agronomic trait. In a still further embodiment of the invention said further agronomic trait comprises resistance and/or tolerance to insects achieved via production of insecticidal proteins such as lectins or proteins derived from *Bacillus thuringiensis*, *Xenorhabdus sp.* and *Photorabdus sp* in the plant. In a still further embodiment of the invention said further agronomic trait comprises resistance and/or tolerance to nematodes achieved via production of nematicidal proteins such as enzyme inhibitors, in the plant

The present invention still further provides a method of selectively controlling weeds in a field said field comprising crop plants and weeds said method comprising applying to said field an agriculturally acceptable formulation of an agrochemical comprising fomesafen and/or acifluorfen wherein the said crop plants are the plants according to the invention.

The present invention still further provides the use of an agrochemical comprising fomesafen and/or acifluorfen to selectively control weeds in a field which field comprises crop plants and weeds comprising applying to said field an agriculturally acceptable formulation of said agrochemical in an amount which is sufficient to be phytotoxic to said weeds but not said crop plants characterised in that said crop plants are the plants according to the invention.

In a further aspect of the present invention there is provided a protein comprising the sequence depicted as SEQ ID No. 10 or a protein variant having a Smith-Waterman score greater than 766 in the SWISSPROT database calculated using the FASTA3 algorithm wherein the said protein variant still encodes a Glutathione-S-transferase. In a further embodiment of the present invention said protein variant has a Smith-Waterman score greater than 770 in the SWISSPROT database calculated using the FASTA3 algorithm wherein the said protein variant still encodes a Glutathione-S-transferase. In a still further embodiment of the present invention said protein variant has a Smith-Waterman score greater than 780 in the SWISSPROT database calculated using the FASTA3 algorithm wherein the

said protein variant still encodes a Glutathione-S-transferase. In a further embodiment of the present invention said protein variant has a Smith-Waterman score greater than 790 in the SWISSPROT database calculated using the FASTA3 algorithm wherein the said protein variant still encodes a Glutathione-S-transferase.

In a further aspect of the present invention there is provided a protein comprising the sequence depicted as SEQ ID No. 7 or a protein variant having a Smith-Waterman score greater than 1094 in the SWISSPROT database calculated using the FASTA3 algorithm wherein the said protein variant still encodes a glutathione-S-transferase.

In a still further aspect of the present invention there is provided a protein comprising the sequence depicted as SEQ ID No. 9 or a protein having a Smith-Waterman score greater than 671 in the SWISSPROT database calculated using the FASTA3 algorithm wherein the said protein variant still encodes a Glutathione-S-transferase. The Fasta algorithm referred to above is well known to the skilled artisan and uses the method of Pearson and Lipman (Proc. Natl. Acad. Sci. USA 85; 2444-2448 (1988)) to search for similarities between one sequence and any group of sequences of the same type as the query sequence. Fasta also determines the best segment of similarity between the query sequence and the sequences in the database , using a variation of the Smith-Waterman algorithm. This "local alignment" procedure is described in Chao, Pearson, and Miller (CABIOS 8; 481-487 (1992)). The score for this alignment is reported as the *opt* and Smith-Waterman score. The database used in the above calculations is the SWISSPROT database. This database, which is well known and frequently used by the person skilled in the art is commercially available from sources such as Geneva Bioinformatics (GeneBio™) S.A 25 avenue de Champel CH - 1206 Geneva Switzerland.

In a further aspect of the present invention there is provided a protein comprising the sequence depicted as SEQ ID No. 1 or a protein variant having a Smith-Waterman score greater than 2152 calculated using the FASTA3 algorithm wherein the said protein variant still encodes a homoglutathione synthetase. In a further embodiment of the present invention said protein variant has a Smith-Waterman score greater than 2159 calculated using the FASTA3 algorithm wherein the said protein variant still encodes a homoglutathione synthetase. . In a still further embodiment of the present invention said protein variant has a Smith-

Waterman score greater than 2169 calculated using the FASTA3 algorithm wherein the said protein variant still encodes a homoglutathione synthetase. . In a still further embodiment of the present invention said protein variant has a Smith-Waterman score greater than 2971 calculated using the FASTA3 algorithm wherein the said protein variant still encodes a homoglutathione synthetase. The opt score can be calculated in accordance with the methods described above. The homoglutathione synthetase referred to above is also capable of catalysing the addition of Beta-alanine onto gamma glutamylcysteine.

The present invention still further provides a method of providing a plant which is tolerant to stress comprising incorporating a DNA encoding a GST or a variant GST according to the present invention into the genome of plant material such that a Glutathione-S-Transferase enzyme is produced wherein said stress is induced by an agent selected from the group consisting of: Ozone; Potassium and Sodium salts; Temperature and Pathogens. Stress tolerant, increased stress tolerance, tolerance to stress and similar phrases used herein refer to the ability of a plant, or group of plants such as a field planted with a particular crop, that are transformed to contain the GST according to the invention as described above, to overcome or resist the effects of a stresser to a greater extent than control-like plants. Increased stress tolerance may vary from a slight increase in the ability to resist or overcome the effects of a stresser to total tolerance where the plant is unaffected by the stresser. Stress tolerant plants may show physical characteristics that indicate increased tolerance to an environmental stress. For example, plants that are transformed to contain the GSTs according to the invention may be larger than control-like plants when grown in the presence of a particular stresser such as hot or cold temperatures. In a particular embodiment the stress tolerant plants according to the invention will be able to tolerate temperatures of at least 1°C higher or lower than the control-like plants. In a further embodiment said plants will be able to tolerate temperatures of at least 2°C higher or lower than the control-like plants. In a still further embodiment said plants will be able to tolerate temperatures of at least 3°C higher or lower than the control-like plants. In a still further embodiment said plants will be able to tolerate temperatures of at least 4°C higher or lower than the control-like plants. In a still further embodiment said plants will be able to tolerate temperatures of at least 5°C or more higher or lower than the

control-like plants. Plants according to the invention may also be able to tolerate any such temperature for a longer time period than that of the control-like plant. Similarly, transformed plants may emerge sooner than non-transformed plants, or be able to grow in the presence of the stresser when the control-like plant is not able to grow under the same conditions. The affected traits of the plant will depend on the type of plant and stresser involved.

The present invention further provides the use of a GST or GST variant according to the invention as described above in a method of producing a plant having increased tolerance to a stress which stress is induced by an agent selected from the group consisting of: Ozone; Potassium and Sodium salts; Temperature and Pathogens. In a particular embodiment of the present invention the GST used in said method encodes the protein depicted as SEQ ID No. 10.

The present invention still further provides the use of a GST or variant GST as described above as a selectable marker gene. When used as a selectable marker the GST according to the invention will normally form part of a transformation construct which comprises a preferred gene of interest along with the selectable marker. When the desired material has been transformed, a selection pressure (created through application of a herbicide such as fomesafen and/or acifluorfen) is applied to the material to initiate and facilitate selection. Only those cells containing the selectable marker survive this step thereby reducing any regenerated material to that which contains the transformation construct. In a particular embodiment said GST comprises a sequence which encodes the amino acid sequence depicted as SEQ ID No. 10 or is a variant GST as described above and selection is achieved through the application of a herbicide which comprises fomesafen and/or acifluorfen at levels which are toxic to control-like material which does not comprise said GST in a similar amount. In a further embodiment of the invention the selection of cells which comprise said marker takes place when the cells have developed into embryogenic callus and/or somatic embryos. Once in possession of the GSTs according to the invention the skilled artisan is quite capable of utilising them as selectable markers.

According to a further aspect of the present invention there is provided a GST protein which is capable of reacting with a monoclonal anitbody raised to the protein depicted as SEQ ID No. 10.

According to a further aspect of the present invention there is provided a GST protein which is capable of reacting with a monoclonal anitbody raised to the protein depicted as SEQ ID No. 1.

The invention will now be described by way of the following non-limiting examples with reference to the figures and sequence listing of which:

Figure 1 shows a schematic diagram of cloning vector pMJB2.

Figure 2 shows a schematic diagram of cloning vector pMOG1051

Figure 3 shows a schematic diagram of binary vector pMOG800

Figure 4 shows a schematic diagram of vector pFT2 used, via *Agrobacterium* transformation, to transform plants to give resistance to herbicides.

Sequence Listing

SEQ ID No. 1. Homoglutathione synthetase from *Glycine max.*

SEQ ID Nos. 2 to 5. Homoglutathione synthetase protein regions.

SEQ ID No. 6. Polynucleotide sequence encoding Homoglutathione synthetase from *Glycine max.*

SEQ ID Nos. 7 to 10. Glutathione-S-transferases 2.6, 3.1, 3.2 and 3.3 respectively.

SEQ ID No. 11 to 14. Polynucleotides encoding the GSTs 2.6, 3.1, 3.2 and 3.3 respectively. (GST 3.3 may also be referred to as GST 3.6).

SEQ ID Nos. 15 to 24. Primers.

SEQ ID Nos. 25 to 34. Soybean sequence P32110 derived nucleic acid sequence; Mungbean sequence U20809; Tobacco sequence Q03663; Potato sequence P32111 derived nucleic acid sequence; Arabidopsis sequence P46421; Arabidopsis sequence P46421 (genomic); Papaya sequence AJ000923; Spruce sequence AF051214; Wheat sequence AF004358; Spruce sequence AF051238 respectively.

SEQ ID No. 35. Primer.

SEQ ID Nos 36 to 43. Sequences of clones listed as SEQ ID Nos. 8, 22, 7, 21, 31, 15, 23 and 25 in International Patent Application Publication Number WO00/18936.

EXAMPLE 1

Isolation and identification of partial cDNAs encoding soybean GSTs.

1.1 Partial cDNAs encoding soybean GSTs were obtained using reverse-transcriptase polymerase chain reaction (RT-PCR) with degenerate oligonucleotide primers designed to regions conserved within known tau-class GSTs using alignment methods known within the art. Total RNA was obtained from soybean cell cultures (*Glycine max* cv. Mandarin) using TRIZOL™ reagent (Life Technologies™) according to the manufacturers guidelines. First-strand cDNA was obtained from 5 µg total RNA using oligonucleotide OG2 (SEQ ID No. 35) in conjunction with Superscript-II reverse transcriptase (Life Technologies™) using standard protocols provided by the manufacturer. Degenerate oligonucleotide primers CON2 (SEQ ID No.17) and CON3 (SEQ ID No.18) were then used in independent PCR reactions with oligonucleotide primer OG9 (SEQ ID No.16) to amplify partial GST-encoding genes from first strand cDNA. 2 µl of the first-strand cDNA synthesised was used as template in the reaction. The PCR conditions used are (94°C, 45 s; 51°C, 30 s and 72°C, 60 s) 35 cycles using Taq DNA polymerase supplied by Life Technologies and a Techne™ thermocycler. The amplified products were ligated into the pCR2.1 vector (Invitrogen™) and transformed into *E. coli* INV F' cells.

1.2 Transformed colonies were selected on LB media containing 100 ug ml-1 ampicillin and 40 ug ml X-GAL. Plasmid was recovered from 5 ml overnight cultures, initiated from each of 90 individual white colonies (45 from each initial RT reaction). The various plasmids were then subjected to restriction analysis (*Eco*R1, *Ssp*1, *Ssp*1: *Sph*1 and *Rsa*1) allowing the grouping of similar clones to be made. Distinct cDNAs were identified by automated DNA sequencing with M13 primers, using an ABI 377 automated DNA sequencer, and analysed for similarity with known GST sequences by conducting database searches using the BLAST algorithm (Altschul S.F *et al.*, 1990).

EXAMPLE 2

Isolation and identification of partial cDNA encoding homoglutathione synthetase.

2.1 A partial length cDNA encoding homoglutathione synthetase was obtained by RT-PCR. Degenerate oligonucleotide primers MS-3 (SEQ ID No.15) and OG9 (SEQ ID No.16) were used to specifically amplify the desired clone from first-strand cDNA, produced from soybean cell cultures as described previously.

2.2 The PCR product obtained was cloned into vector pCR2.1 and sequenced using an ABI377™ automated DNA sequencer. Products exhibiting homology to related glutathione synthetases were identified by searching the databases using the BLAST algorithm.

EXAMPLE 3

Library construction, clone isolation and sequencing.

3.1 A cDNA library was constructed using the lambda ZAP-II system (Stratagene). Total RNA was isolated from 5-day-old soybean cell suspension cultures (cv Mandarin) using TRIZOL™ reagent. Poly A+ mRNA is isolated from total RNA using PolyATtract (Promega™) according to standard supplied protocols. The cDNA library was constructed following standard protocols provided by Stratagene™.

Table 1 below shows the characteristics of soybean cDNA library constructed from 5-day old soybean cell suspension cultures (cv. Mandarin).

Library	Primary titre	Average cDNA size
Cell culture seedling	$2x10^6$ pfu	1.4 Kb

3.2 The partial cDNA sequences identified in previous examples were ^{32}P-labelled using a Ready-to-go labelling kit (Pharmacia™) and used to screen the soybean cDNA library for full-length cDNAs. 160,000 pfu's were screened and putative colonies cross-hybridising with the probes subjected to secondary and tertiary screening until plaque purity was observed. Plasmid DNA was recovered from the plaque pure stocks using in vivo excision protocols provided by

Stratagene™. Full-length cDNAs encoding GSTs and a full length cDNA encoding the homoglutathione synthetase was identified

EXAMPLE 4

Identification of soybean GSTs with activity toward herbicide substrates.

4.1 Bacterial expression of cDNA clones

Full length cDNAs encoding homoglutathione synthetase and glutathione *S*-transferase were independently expressed in *E.coli* using the pET expression system (Novagen™).

4.2 *Nde* I or *Nco* I sites were introduced as appropriate into the 5' end cDNAs encoding glutathione *S*-transferase and *Bam* H1 at the 3' end using PCR. The cDNA was then cloned into pET-24a or pET-24d as appropriate. The resulting plasmids were introduced into *E. coli* BL21 (DE3) using standard bacterial transformation procedures known to the skilled artisan. Expression and purification of the recombinant GST was performed using anion-exchange chromatography and *S*-Hexyl-glutathione affinity chromatography according to the methods known in the art (for example, Skipsey *et al.*, 1997). GST activity of the purified recombinant protein toward CDNB, acetochlor, acifluorfen, chlorimuron-ethyl, fluorodifen. fomesafen and metolachlor in the presence of both glutathione and homoglutathione was performed via assays which are described in Andrews *et al.*, 1997 and Skipsey *et al.*, 1997.

4.3 Table 2 below shows activity of recombinant enzymes. ±SE, n=2, ND = Not detectable, * Activity nkat mg-1, ** Activity pkat mg-1

Table 2

Clone	GSH	hGSH	GSH	hGSH	GSH	hGSH	GSH	hGSH	GSH	hGSH	GSH	hGSH	GSH	hGSH
GST 2.6	18.7 ±0	19.0 ±0	352.5 ±106.5	391.0 ±12.0	ND	31.1 ±0	ND	ND	ND	19.3 ±2.8	27.1 ±1.4	350.5 ±78	142.0 ±15.0	95.5 ±12.5
GST 3.1	59.2 ±0	59.2 ±0	190.5 ±13.5	11.0 ±2.5	152.5 ±1.5	238.5 ±2.5	ND	ND	1150 ±50	1350 ±0	213.6 ±2.6	484.6 ±11.8	177.5 ±7.5	67.5 ±1.5
GST 3.2	24.9 ±2.3	10.3 ±1.6	418.0 ±15.0	143.0 ±18.0	7.5 ±0.4	23.2 ±0.1	ND	ND	78.2 ±1.4	68.0 ±4.3	21.3 ±7.0	201.6 ±2.7	253.0 ±56.0	191.5 ±1.5
GST 3.3	82.9 ±4.2	189.0 ±9.7	420.0 ±30.0	ND	123.0 ±2.0	1350 ±31	ND	ND	360.5 ±48.5	602.5 ±100.5	110 ±3.0	6250 ±150	312.0 ±3.0	264.5 ±8.5

EXAMPLE 5

Identification of active Homoglutathione synthetase.

5.1 Bacterial expression of cDNA clone.

A 5' *Nco* 1 and 3' *Xho* 1 restriction enzyme site were introduced into the homoglutathione synthetase cDNA via PCR using primers MS4-*Nco* (SEQ ID No.19) and MS-4-HIS (SEQ ID No.20). Additionally, the use of the MS-4-HIS primer provided a C-terminal HIS-tag to assist purification of the enzyme.

5.2 The resulting PCR fragment was digested with *Nco*1 and *Xho* 1 and ligated into similarly digested pET-24d. This vector was termed pET-MS4-His.The purification of homoglutathione synthetase was performed using the following method. *E. coli* BL21 (DE3) harbouring the pET-MS4-His plasmid (section 1.3) were grown at 30 °C until OD_{600}=0.5, after which isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 0.1 mM. Following a 3 hour incubation the bacteria were collected by centrifugation, re-suspended in buffer A (20mM Tris, 0.5 M NaCl, 5 mM imidazole pH 8.0) and lysed. Cell debris was removed by centrifugation (10,000 g, 10 min) and the supernatant applied to a 5 ml iminidiaceic acid column (Sigma™), previously charged with $NiSO_4$ and equilibrated in buffer A. The column was washed with buffer A containing 20 mM imidiazole, followed by buffer A containing 300 mM imidiazole to remove the affinity bound protein.

5.3 The eluted protein was concentrated using a Centriplus 30 (Amicon™) spin column and re-suspended in buffer A prior to application onto a 1 ml HiTrap chelating column (Phamacia™), pre-charged with $NiSO_4$ and equilibrated as described previously. Affinity bound protein was then recovered using an increasing concentration of 20-200 mM imidiazole and the presence of His-tagged recombinant homoglutathione synthetase was detected using His-tagged antibodies according to known procedures. Fractions containing recombinant homoglutathione synthetase were pooled, concentrated using Centricon 30 spin columns and re-suspended in 20 mM Tris-HCl pH 8.0, 1mM DTT.

5.4 Homoglutathione synthetase was assayed for activity in 250 mM Tris-HCl pH 8.0, 50 mM KCl, 20 mM MgCl2, 5 mM DTT, 10 mM ATP, 1 mM γ-glutamylcysteine and 50 mM glycine or 10 mM β-alanine in a total 100 µl. Experimental controls did not contain the

enzyme or glycine/β-alanine. Assays were performed at 30°C for 60 min, with 20 µl aliquots removed at regular time intervals. Monobromobimane derivatisation was then performed on the aliquot to determine the presence of either glutathione or homoglutathione in accordance with the methods described by Cummins *et al.*, 1997.

Table 3 - Characteristics of recombinant homoglutathione synthetase.

	β-alanine	Glycine
Vmax (pkats/mg hGS)	42250 ± 3173	3300 ± 341
Km (mM)	0.84 ± 0.19	54.9 ± 18.7

EXAMPLE 6

Generation of herbicide resistant/tolerant plants

6.1 Plants with increased tolerance to the herbicides acifluorfen and fomesafen were obtained by expression in the plant host of both the active soybean GST 3.3 and the homoglutathione synthetase to further increase levels of homoglutathione, required for the enhanced efficiency function of GST 3.3.

6.2 PCR, using oligonucleotide primers hGSH-Nco1 (SEQ ID No. 21) and hGSH-Kpn1 (SEQ ID No.22) was used to introduce a 5' *Nco*1 site and 3' *Kpn*1 site into the GST 3.6 cDNA. The resulting PCR product was purified, sequenced, digested with *Nco*1 and *Kpn*1 and ligated into the vector pMJB2 (Figure 1). The expression cassette, comprising the double enhanced CaMV35S promoter : Glucanase II leader, hGSH synthetase cDNA and *nos* terminator was then excised from pMJB2 using *Hind* III/ *Eco*R1 and ligated into the similarly digested binary vector pMOG800 (Figure 3).

6.3 Oligonucleotide primers 3.6-Bgl II (SEQ ID No.23) and 3.6 *Nco* 1 (SEQ ID No.24) were used to introduce *Nco*1 and *Bgl* II sites at the 5' and 3' end of the GST 3.6 cDNA respectively. The resulting PCR product was purified, sequenced, digested with *Nco*1 and *Bgl* II and sub-cloned into the vector pMOG1051 (Figure 2). The expression cassette, comprising the RolD/Fd promoter, GST 3.3 cDNA and potato PI-II terminator was then excised from pMOG1051 using *Bam* H1 and ligated into the unique *Bam* H1 site in the binary vector pMOG800 (Figure 3) harbouring the homoglutathione synthetase expression

cassette (see above). Orientation of the insert was determined by PCR. The resulting binary vector, termed pFT2 (Figure 4) was sequenced in entirety to confirm authenticity.

EXAMPLE 7

Plant Transformation and Regeneration.

7.1 The binary vector pFT2 was transformed into *Agrobacterium tumefaciens* strain LBA 4404 using the freeze thaw method of transformation provided by Holsters *et al.*, 1978. Tobacco transformation and whole plant regeneration was performed using *Nicotiana tabacum* var. Samsun according to standard protocols detailed by Bevan, 1984. Transformation events were selected on MS-media containing kanamycin. Other transformation methods may be employed for the transformation of other crops, such as soybean. Examples of such methods are referred to below.

EXAMPLE 8

Analysis of Transgenic Plants

8.1 PCR analysis of transformants

Leaf samples were taken from transformed lines and DNA extracted according to the methods described by Edwards et al., 1991. Oligonucleotide primer pairs SEQ ID No. 21 and SEQ ID No.22 and SEQ ID No.23 SEQ ID No.24 were used to amplify specific regions within the homoglutathione synthetase and GST 3.3 to enable the detection of both the transgenes within the putative transformed plant material.

The following analyses were performed on PCR positive material in order to confirm the functionality of the introduced genes.

8.2 Enzyme Analysis

GST activity toward fomesafen in transgenic lines was performed in accordance with the methods described in the publication of Andrews *et al.*, 1997.

8.3 Free thiol synthetase activity

Glutathione and homoglutathione synthetase activities were determined in the transgenic lines using the assay described previously.

8.4 Free thiol determination

The presence of free thiol within the transgenic plants was determined using the method of monobromobimane derivatisation. Tissue (1g) is weighed accurately, frozen in liquid nitrogen, ground to a fine powder and transferred to a clean tube containing 3 ml 0.1 M HCl. After incubation on ice for 30 min with occasional mixing, the slurry was transferred to an eppendorf tube and centrifuged (13,000 g, 3 min).Two 100 µl aliquots of supernatant were transferred to clean eppendorfs, and 10 µl water added to one and the other spiked with 1 mM glutathione or homoglutathione (10 µl). The free thiol present was reduced by adding 10 µl 1 M NaOH followed by 10 µl 1 M NaOH containing 20 mg ml $NaBH_4$ and the solution incubated for 10 min at room temperature.

8.4.1 The reaction was stopped by the addition of 120 µl 3.6 M HCl and the samples centrifuged (13,000 g, 5 min).The supernatant (100 µl) was transferred to a fresh tube and 10 µl 5 mg ml monobromobimane dissolved in acetonitrile added, followed by 5 µl of 35% v/v N-ethylmorpholine. The samples were placed in the dark for a period of 20 min, and the reaction stopped by the addition of 880 µl 5% acetic acid (5 ul v/v). A standard curve was prepared by derivatising glutathione or homoglutathione (0-20 mmol), and the *S*-bimane conjugates analysed by HPLC using methodology known in the art (Cummins *et al.*, 1997).

EXAMPLE 9

Herbicide tolerance tests

9.1 Following tissue culture, kanamycin-resistant plantlets were transferred to 5 inch pots containing John Innes potting compost no. 3. The plants were allowed to develop to approximately the 10-leaf stage and fomesafen applied at 10g ai ha^{-1}, formulated with non-ionic surfactant, to the aerial tissue using a track sprayer. Visual assessment of phytotoxicity/plant necrosis is performed 5 days post application.

Table 4 Analysis of transgenic tobacco lines.

Plant line	% damage	GST activity towards fomesafen		GSH activity	hGSH activity	ug GSH	UG hGSH
		GSH	hGSH				
WT	55	1.4	5.0	0.91	ND	71.3	ND
T82 3740	15	5.2	558.5	4.6	113.3	78.3	18.5
T82 3747	5	10.7	1320.4	1.59	145.9	112.2	26.9

EXAMPLE 10

Production of homozygous plant lines

10.1 Single copy transgenic plant lines were identified by Southern blot analysis according to methods described by Sambrook, 1989 using appropriate radiolabelled probes. Segregation analysis was performed on plants containing single insertion events by germination on MS media containing kanamycin. Further confirmation of homozygous lines may be performed by back crossing transgenic lines with wild-type tobacco and analysis genetic segregation following selection on kanamycin.

References

Altschul S.F., Gish W., Miller W., Myers E.U and Lipman D.J (1990). Basic local alignment tool. *Journal of Molecular Biology* **215** 403-410.

Andrews C.J., Skipsey M., Townson J.K., Morris C., Jepson I. and Edwards R (1997) Glutathione transferase activities toward herbicides used selectively in soybean. *Pesticide Science* **51** 213-222.

Bevan M.M (1984) Binary *Agrobacterium* vectors used for plant transformation. *Nucleic Acids Research* **12** 8711-8721.

Cummins I., Moss S., Cole D.J and Edwards R (1997) Glutathione Transferases in Herbicide-resistant and Herbicide-Susceptible Black-grass (*Alopecurus myosuroides*). *Pesticide Science* **51** 244-250.

Edwards K., Johnstone C and Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19** 1349.

Holsters M., de Waele D., Depicker A., Messens E., van Montagu M and Schell J (1978) Transfection and transformation of *Agrobacterium tumefaciens. Molecular and General Genetics* **163** 181-187.

Skipsey M., Andrews C.J., Townson J.K., Jepson I and Edwards R (1997) Substrate and thiol specificity of a stress-inducible glutathione transferase from soybean. *FEBS Letters* **409** 370-374.

Christou, Paul. Physiol. Plant. (1990), 79(1), 210-12. Soybean transformation by electric discharge particle acceleration.

Dan, Yinghui; Reichert, Nancy A. (1999), US Patent No. 5968830 Soybean transformation and regeneration methods.

Williams, Edward J.; Emler, Carol A.; Julson, Lori S.; Martinell, Brian J.; Mccabe, Dennis E.; Huang, Yong. (2000) International Patent Application Publication Number. WO/0042207 Soybean transformation method omitting callus culture.

Hinchee, Maud Ann Wrightson; Clemente, Thomas Elmo; Connor-Ward, Dannette Vaudrilyn; Fedele, Mary Jacqueline; Fry, Joyce Ellen; Howe, Arlene R.; Rozman, Renee Jean. US Patent No. 5959179. Method for transforming soybeans.